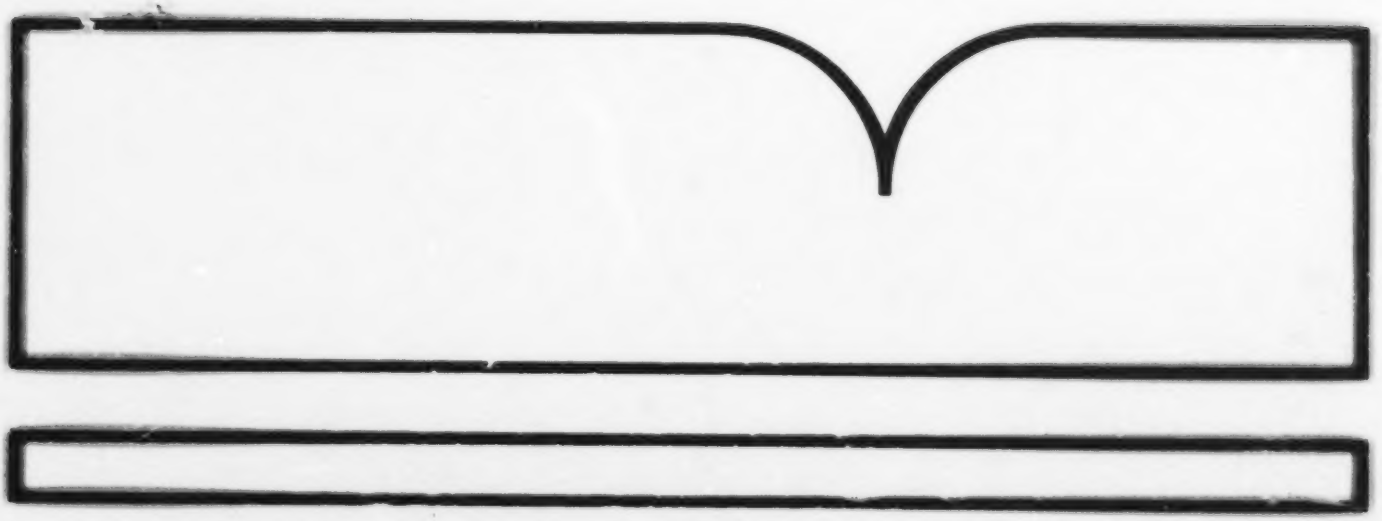
PB89-201834

Modeling Ambient Carbon Monoxide Trends to Evaluate Mobile Source Emissions Reductions

(U.S.) Environmental Protection Agency
Research Triangle Park, NC

1987



U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA/600/J-87/477	2.
3. RECIPIENT'S ACCESSION NO. PB89 201834/AS	
4. TITLE AND SUBTITLE Modeling Ambient Carbon Monoxide Trends to Evaluate Mobile Source Emissions Reductions	5. REPORT DATE
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Robin L. Dennis and Mary W. Downton	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Robin L. Dennis, Atmospheric Sciences Research Laboratory--RTP, NC, U.S. EPA, RTP, NC 27711 Mary W. Downton, National Center for Atmospheric Research, Boulder, CO 80307	10. PROGRAM ELEMENT NO. CDTA1D / 04-1315 (FY86)
	11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Atmospheric Sciences Research Laboratory--RTP, NC Office of Research and Development U. S. Environmental Protection Agency Research Triangle Park, NC 27711	13. TYPE OF REPORT AND PERIOD COVERED
	14. SPONSORING AGENCY CODE EPA/600/09

15. SUPPLEMENTARY NOTES

Journal of Climate & Applied Meteorology, Vol. 26, No. 10, Oct. 1987, pp. 1376-1391

16. ABSTRACT

Regression models have been used with poor success to detect the effect of emission control programs in ambient concentration measurements of carbon monoxide. An advanced CO regression model is developed based on an understanding of the physical processes of dispersion. Its performance is shown to be superior to the more traditionally developed regression and time series models. The model reliably separates the effects of emissions change from the effects of fluctuations in meteorological conditions. The model has sufficient precision to identify present trends in emissions from ambient concentration data. This model should be useful for detecting changes in emission trends due to implementation of a control program on vehicular emissions such as an inspection and maintenance program.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

18. DISTRIBUTION STATEMENT RELEASE TO PUBLIC	19. SECURITY CLASS (This Report) UNCLASSIFIED	21. NO. OF PAGES 19
	20. SECURITY CLASS (This page) UNCLASSIFIED	22. PRICE

EPA/600/J-87/477
PB89-201834

Modeling Ambient Carbon Monoxide Trends to Evaluate Mobile Source Emissions Reductions

ROBIN L. DENNIS

MARY W. DOWNTON

Reprinted from JOURNAL OF CLIMATE AND APPLIED METEOROLOGY, Vol. 26, No. 10, October 1987
American Meteorological Society

REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL INFORMATION SERVICE
SPRINGFIELD, VA. 22161

Reprinted from JOURNAL OF CLIMATE AND APPLIED METEOROLOGY, Vol. 26, No. 10, October 1987
American Meteorological Society

Modeling Ambient Carbon Monoxide Trends to Evaluate Mobile Source Emissions Reductions†

ROBIN L. DENNIS

Atmospheric Science Research Laboratory, U.S. EPA, Research Triangle Park, NC 27711

MARY W. DOWNTON

*The National Center for Atmospheric Research, Boulder, CO 80307**

(Manuscript received 31 May 1986, in final form 10 April 1987)

ABSTRACT

Regression models have been used with poor success to detect the effect of emission control programs in ambient concentration measurements of carbon monoxide. An advanced CO regression model is developed whose form is based on an understanding of the physical processes of dispersion. Its performance is shown to be superior to the more traditionally developed regression and time series models. The model separates the effects of emissions change from the effects of fluctuations in meteorological conditions. The separation appears to be most reliable for winter conditions. The model has sufficient precision to identify present trends in emissions from ambient concentration data. This model should be useful for detecting changes in emission trends due to implementation of a control program on vehicular emissions such as an inspection and maintenance program.

1. Introduction

Under the U.S. Clean Air Act Amendments of 1977, urban areas which were expected to continue to exceed ambient standards for carbon monoxide (CO) and/or ozone by the end of 1982 were required to adopt an inspection and maintenance (IM) program for light duty motor vehicles. The purpose of an IM program is to correct problems of equipment adjustment and tampering which erode the effectiveness of the federal tailpipe emissions standards. The IM programs are implemented by the states, but must pass review by the U.S. Environmental Protection Agency.

Renewal of the IM program often hinges on demonstrating that the IM program is effective. Two approaches to such a demonstration are possible: studies of vehicle emissions before and after IM, and studies of ambient concentrations before and after. Vehicle emission studies are more likely to discern an effect due to IM upon vehicle emissions than are ambient concentration studies. However, emissions studies are considered by some to be unrepresentative of the actual changes in emissions of the vehicle fleet, because tampering will not be accounted for accurately. Legislatures in a number of states have asked that effectiveness of IM be shown in terms of reductions in ambient concentrations. The ability to assess IM effectiveness through studies of ambient concentrations is therefore important.

A criterion for effectiveness based on ambient concentration studies is very difficult to meet. One cannot simply establish the trend in CO emissions from direct measurements of CO concentrations. One must estimate average reductions in emissions per vehicle while total vehicular activity may rapidly be increasing. Most importantly, ambient CO concentrations fluctuate greatly from day to day and year to year as a result of changing meteorological conditions. To detect the effects of emission control programs requires a stochastic model of ambient CO which is capable of reliably separating the effects of emissions change from fluctuations in meteorological conditions and trends in traffic patterns.

Time series and regression models have been used by Dr. Tiao and others for analyses of this type (e.g., Tiao et al., 1982; Ledolter and Tiao, 1979; Tiao et al., 1975). Generally, improvements in CO concentrations associated with reductions in emissions have been found. However, it has been difficult to separate the effect of IM programs from other changes, such as the increasingly strict federal emission standards for new cars. Frequently, the reductions attributable to IM have not been statistically significant at a 95% confidence level. More accurate and more precise time series and regression models are needed.

† Although the research described in this article has been supported in part by the United States Environmental Protection Agency, it has not been subjected to agency review and therefore does not necessarily reflect the views of the agency and no official endorsement should be inferred.

* NCAR is sponsored by the National Science Foundation.

© 1987 American Meteorological Society

NTIS is authorized to reproduce and sell this report. Permission for further reproduction must be obtained from the copyright owner.

One possible reason for insufficient precision is that the models used by Tiao and his associates were developed empirically, by examination of data, rather than from physical principles. The possible model forms were often limited by unavailability of both long-term local traffic data and potentially important meteorological measurements. More completely correct models based on physical theory and adequate data could both increase the precision of the parameter estimates and assure accurate interpretations of their meaning.

In this paper, we develop a regression model for ambient CO concentrations by looking first at the equational forms of physical models based on meteorological principles. Three methods of simplifying the physical models are considered, in order to suggest useful regression model forms and data needed to support estimation of the statistical models. Then several regression models, based on the form of the simplified physical equations, are evaluated using Denver, Colorado data. A significant improvement in precision is achieved with the best regression model, compared to past model forms. The performance of the best regression model is compared to the emissions reductions predicted for Denver by an "outside expert", the U.S. Environmental Protection Agency's computer program MOBILE2.5, which is based on direct measurements of vehicle emissions, auto fleet characteristics, and driving patterns for Denver. The performance of this regression model is also compared with that of a time series model similar to those used by other researchers to analyze ambient CO data to estimate trends in vehicular emissions for other cities. Then the sensitivity of this model to detecting changes in emissions trends is assessed. Finally, because greater improvements are still desirable, some improvements in the best regression model are attempted.

2. Physical basis

Regression models must be relatively simple in order to minimize the number of parameters to be estimated. In using a simplified model, interpretation of fitted parameters is complicated by the possible impact of variables that have been omitted from the model. If the omitted variables have a purely random effect, then the precision of the parameter estimates may be reduced and the residual error term of the model will be enlarged. A more serious problem occurs when the omitted variables have nonrandom effects, leading to biased parameter estimates and possibly an inaccurate physical interpretation of the resulting model or model parameter estimates. To evaluate the effectiveness of emission controls, it is important that the parameter in the model which represents vehicular emissions be as little distorted as possible by other influences.

We believe it is important to constrain the form of regression models to be similar to a simplification of a more exact description of air pollution dispersion. Starting with physical theory ensures a proper identification of the important processes and their relationship. To begin, we take two different approaches to identification of model form. The first is a box-model approach to the conservation of mass equation and the second is a scale analysis of that same equation.

Box-model style arguments involve simplifying the turbulent diffusion equation by giving up spatial resolution to achieve an analytical solution. The emphasis is placed on explaining temporal variation of the pollutant concentrations. One such solution for concentrations of nonreacting pollutants in an urban area was proposed by Lettau (1970). Assuming no flux of pollutants through the top of the box, the simplified equation is

$$\frac{\partial c}{\partial t} = \frac{Q}{h} - c\frac{U}{X} \tag{1}$$

where

Q the ground-level area source emissions rate,
U the bulk wind,
h the mixing depth,
X the diameter of the area of ground emissions,
c bulk pollutant concentration.

Here, Q is a function only of time t, and c is the bulk value of the pollutant concentration in the volume of city air. The time-dependent solution is

$$c_t = c_{t0}\exp\left(-\frac{U}{X}t\right) + \frac{QX}{Uh} - \frac{QX}{Uh}\exp\left(-\frac{U}{X}t\right) \tag{2}$$

where c_{t0} is the initial concentration at time $t = 0$. The steady state solution is

$$c_{ss} = \frac{QX}{Uh} \tag{3}$$

where c_{ss} is the bulk concentration in the box of height h and width and length X. Equation (3) is the simplest representation of the dominant meteorological processes of emissions, advection and mixing depth, describing how they influence the pollutant concentration. The diameter of the area of emissions is treated as a constant in this application.

A second method, scale analysis, can also be used to ascertain the importance of the different processes affecting pollutant concentrations (Anthes, 1979). To carry out a scale analysis, we define the urban domain such that the characteristic scale of variation of the pollutant concentration is comparable to the size of the domain itself. We then define dimensionless variables for the mass conservation equation as follows:

$$c' = c/C,$$
$$u' = u/U,$$
$$w' = w/W,$$

$$x' = x/X,$$
$$y' = y/X,$$
$$z' = z/h,$$
$$t' = t/T,$$

where C is the mean amplitude of variation of pollutant concentration, U and W are the magnitudes of the mean horizontal and vertical velocity, X is a typical horizontal scale associated with the spatial variation of C, h is the vertical scale of variation, and T is the time scale of variation of C.

Assuming K theory, the mass conservation equation may now be written as

$$\frac{1}{T}\left(\frac{\partial c'}{\partial t'}\right) = -\frac{U}{X}\left(u'\frac{\partial c'}{\partial x'}\right) - \frac{W}{h}\left(w'\frac{\partial c'}{\partial z'}\right) + \frac{K_H}{X^2}\left(\frac{\partial^2 c'}{\partial x'^2}\right) + \frac{K_H}{X^2}\left(\frac{\partial^2 c'}{\partial y'^2}\right) + \frac{K_z}{h^2}\left(\frac{\partial^2 c'}{\partial z'^2}\right) + \frac{Q}{Ch} \quad (4)$$

where K_H and K_z are the horizontal and vertical diffusivities, respectively. By definition, all terms in parentheses are dimensionless and of order of magnitude 1. Therefore, the relative importance of the different processes is given by the coefficients of the dimensionless terms.

Noting that the sign of the second partials is negative, the quasi-steady state concentration is

$$-\frac{U}{X} - \frac{W}{h} - \frac{K_H}{X^2} - \frac{K_z}{h^2} + \frac{Q}{Ch} = 0 \quad (5)$$

where U/X is associated with advection, W/h with subsidence or inward flux from the top, K_H/X^2 with horizontal mixing, K_z/h^2 with vertical mixing and Q is the source term. The second and third terms are small relative to the others. Therefore, scale analysis indicates the following model form:

$$C_{qss} = \frac{Q/h}{(1/X)U + (1/h^2)K_z}. \quad (6)$$

The horizontal extent of the region, X, is taken to be a constant here, as well.

Both the simplified box-model approach and the scale analysis indicate that three major processes, namely, emissions, advection and vertical dispersion contribute to the time variation of the pollutant concentration levels and should be included in any model form. One should not assume that advection can be substituted for vertical dispersion in a model of the diurnal behavior of pollutants, as has been done by Tiao and other statisticians. Such substitution may thwart an attempt to obtain the most precise fit possible. Advection and dispersion do not follow precisely the same diurnal pattern: Vertical dispersion follows the heat flux which peaks shortly after midday. Wind speed, however, because of heat storage at the surface and continued afternoon mixing, tends to peak later in time than the heat flux; or the wind speed reaches a maximum about midday and stays on a plateau until mid- to late afternoon. The distinction between vertical diffusion and wind speed is expected to be reduced if the solar heating is not strong due to significant overcast.

The form a statistical model should take is also indicated by the previous analyses. The box-model solution suggests a multiplicative interaction between wind and diffusivity, while scale analysis suggests an additive interaction between them. One advantage of the additive form is that the equation does not go to infinity with zero wind speed. Both forms are investigated here and show no difference in performance.

One should not stop with steady state model forms, however. Equation (2) indicates that for a 1 h averaging time, a wind speed of 5 m s^{-1} (18 km h^{-1}) and a 20 km distance of influence of emissions, the exponential term is approximately $1/e$. Significant departures from steady state can occur routinely because vehicular emissions and the meteorological variables (U and h) could (and often do) change rapidly with respect to time scales of an hour. Thus, we need to extend the simplified box-model form to account for the non-steady state and highly time-dependent (serially correlated) character of pollutants associated with vehicular emissions.

For this extension, we draw upon the work of Long et al. (1983). They have developed an improved simplified box model of pollutant concentration averaged over a specified averaging period for the prediction of a time series of concentrations. From the mass conservation equation, they derive an equation nearly identical to Eq. (2). The pair $\overline{Ch}$ is solved instead of C alone, because $\overline{Ch}$ is not necessarily equal to $\bar{C}\,\bar{h}$ since both C and h are time dependent. Integrating the equation for an averaging period, T, results in the following solution:

$$\overline{Ch} = \frac{E}{XU}\left\{1 - \frac{x}{TU}\left[1 - \exp\left(-\frac{U}{X}T\right)\right]\left[1 - \frac{C_0 h_0}{E/(XU)}\right]\right\} \quad (7)$$

where C is the pollutant concentration, h the mixing height, E the total urban emission rate (mass per time T), U a bulk wind speed throughout the box, and X is the width and length of the urban box.

Assuming a characteristic mixing height, $h^* = \overline{Ch}/\bar{C}$ can be defined during the averaging period T; one obtains the final model form:

$$C = \frac{E}{XUh^*}\left\{1 - \frac{x}{TU}\left[1 - \exp\left(-\frac{U}{X}T\right)\right]\left[1 - \frac{C_0 h_0}{E/(XU)}\right]\right\}. \quad (8)$$

The specification of the current hour's initial conditions, $C_0 h_0$, is critical in predicting the time-averaged concentration for the current hour. It has been found

that C_0 and h_0 can reliably be equated with the previous hour's conditions, C_{t-1} and h_{t-1} (Long et al., 1982). The last term in Eq. (8), $\{1 - C_0 h_0[E/(XU)]^{-1}\}$, has the quality of a lag term. This term can be either positive or negative, depending on whether last hour's concentration times effective mixing height is greater or less than this hour's steady state concentration times effective mixing height. This changing of the sign is a major difference from traditional approaches to incorporating lags in regression models. A regression model based on Eq. (8) outperformed the other regression models and was extensively used in this analysis.

3. Data used in the regression models

The physical models discussed earlier require data related to emissions, advection, vertical dispersion, and pollutant concentration. Data routinely collected in the Denver metropolitan area include hourly CO concentrations, traffic counts, and wind speed and direction. We are interested in determining the change in average emission rate over the years; therefore, emissions can be modeled as emission rate (grams/car over a unit distance) times the amount of traffic passing near the pollution monitor. If the amount of traffic were increasing, then one would need to correct for the increase in emissions per vehicle as speed decreases. However, at the locations we studied, local traffic counts were relatively constant; thus we could ignore vehicle speed and rely solely on traffic count. In this analysis, wind speed is used to measure advection. Wind direction is ignored because of the urban location of the pollution monitors, which are essentially surrounded by busy streets. Unfortunately, no direct measure of vertical dispersion exists for Denver (and most places). Therefore, an empirical measure of vertical dispersion for Denver was developed as part of this study.

Several independent datasets were used in order to cross-check the consistency and interpretability of the regression models. Two of the seven monitoring sites in the Denver region, CAMP and Arvada, were selected because nearby (within a few blocks) hourly traffic counts were available from which to estimate local emissions and no changes had been made to the traffic counters, the monitors or in monitoring protocol during the 1976–80 period of this study. The CAMP station is downtown in a high traffic area and usually has the highest CO concentrations in the region. The Arvada station is in a suburb west of Denver, with CO concentrations similar to those at other suburban stations and about two-thirds of those at CAMP.

The data from these two sites were divided into separate datasets for summer and winter, weekdays and weekends, because daily patterns of CO concentrations (responding to emissions and meteorology) are highly dependent on season and day of week of the original hourly measurements. The data used in the models are in the form of average diurnal patterns, i.e., a seasonally averaged value for each hour of the day, for each of the 5 years, 1976–80. The time period of 1976–80 does not include an inspection and maintenance (IM) program in Denver, which greatly simplifies the evaluation of the models. The Denver IM program did not take effect until 1982. Separate seasonally averaged datasets were established for winter weekdays, winter weekends, and summer weekdays at each of the two monitoring sites. (Summer weekend data were not used because CO concentrations were consistently very low throughout the day with little variation.) This produced six independent datasets on which to test the models.

Graphs of the diurnal patterns of hourly CO, hourly traffic count, and hourly wind speed for winter 1979 are shown in Fig. 1. The shapes of the patterns are similar for other years. The lag in wind speed relative to the heat flux cycle is quite apparent. Summer wind speeds show the same lag relative to the overall heat flux cycle, peaking later in the day. Summer traffic patterns are quite similar to the winter ones, while summer CO concentrations are lower and have a much smaller afternoon peak. The CO and traffic patterns were compiled separately for weekdays and weekends. Mean wind speeds from the small weekend samples proved to be highly unstable, however, so wind speeds were averaged over all days of the week. This instability is not surprising because the distribution of hourly wind speeds is highly skewed and a single extreme value has a large impact on the mean of a small sample. Since meteorological conditions do not differentiate systematically between weekdays and weekends, the use of mean wind speeds for all days of the week is not expected to bias the analysis.

- *Estimate of vertical dispersion.* Routine monitoring in urban areas does not include a direct measure of vertical dispersion. Tiao et al. (1982) attempted to use surrogate measures, such as the inverse of relative humidity, but such models proved to be little better than those which used only wind speed. We believe that relative humidity (RH) is an inappropriate surrogate because it responds to different meteorological influences than does dispersion. The interannual variation of 1/RH is generally different from that of vertical dispersion. In preliminary regression models of the Denver data, wind speed was clearly not an adequate surrogate for all forms of dilution and advection. This inadequacy is not surprising because the minimum CO concentrations are at the time of the maximum in the solar heating cycle, yet the peak in the diurnal cycle of wind speed is later in time than the peak in the solar heating cycle (see Fig. 1). Incorporation of a diurnal pattern of vertical dispersion appeared to be essential for Denver, which has good solar heating.

For a first estimate, we considered the hour-by-hour depths through which there is considerable mixing (effective mixing depth), as distinct from mixing heights,

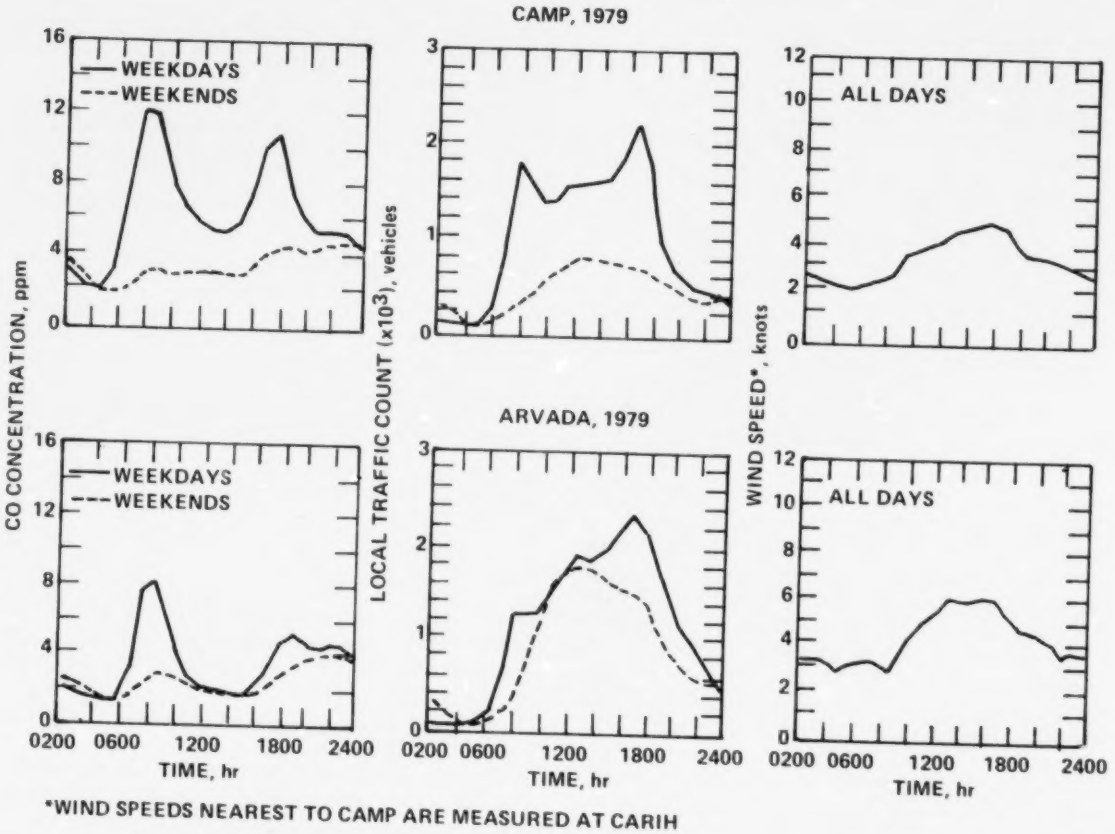


FIG. 1. Average diurnal patterns of carbon monoxide (CO), traffic counts, and wind speed at CAMP and Arvada stations for winter, 1979. Windspeeds nearest to CAMP are measured at CARIH.

based on relative vertical diffusivities at the surface. This approach neglects the fact that at night the effective mixing height is determined by mechanical turbulence, rather than thermal gradients. Vertical diffusivities vary over a diurnal cycle by a factor of 10 to 100. For the regression model, it is sufficient to express the dispersion in relative terms. An exponent on the variable will allow the model to empirically adjust the maximum to minimum ratio as part of the fitting. As a start, we chose a ratio of 100.

The diurnal shape was defined by considering the evolution of the convectively driven mixed layer (Carson, 1973). The early morning increase in vertical mixing was assumed to linearly increase by a factor of 10 between sunrise and the time of inversion fill-in. The time to inversion fill-in was derived for several representative cases using the model of Tennekes (1973). After inversion fill-in at midmorning, a rapid rise occurs with vertical mixing increasing by another factor of 10. The rapid rise peaks at noon and remains until the surface temperature passes its peak. The fall of the effective mixing depth in the evening was slowed to account for the heat island effects typical in urban areas. The diurnal pattern of the diffusion index, D, representing the effective mixing depth which provides the best fit in the regression models for CAMP summer and winter is shown in Fig. 2. This shape is nearly identical to the diurnal variation of mixing depths observed for St. Louis which indicate a maximum to minimum ratio of 10 (Godowitch, personal communication). The maximum-to-minimum ratio giving the best fit for the two Denver sites was much less than that expected from vertical diffusivity arguments, being 2.5 for the downtown site and 5 for the suburban site.

For the second estimate of vertical dispersion, we used the simplified box model of Long [Eq. (8)] to empirically extract an estimate of the effective mixing height. A representative month of CO data was used with the box model for extracting h^*. The diurnal variation of h^* was cross-checked using several years of coefficient of haze data at CAMP and Arvada. The predictions compared well with the observations, given the uncertainties in the emissions inventory. The shape of this vertical dispersion term produced consistently

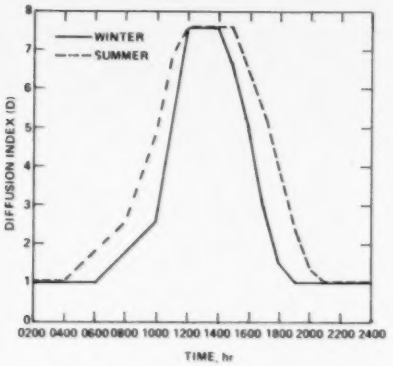


FIG. 2. First diurnal pattern of relative vertical mixing used with early regression models.

biased residuals for certain morning and evening hours. In the morning, it probably did not account for the accelerated ventilation due to the entrainment of the larger momentum in the nocturnal jet at the top of the mixed layer. During this period, the surface wind speed is not sufficiently representative of the ventilating winds throughout the mixed layer. We adjusted the shape slightly to reduce the bias. The shapes that gave the best fit for CAMP and Arvada are shown in Fig. 3.

Experience with the different regression models indicated that the first vertical dispersion estimate performed well with model forms based on Eq. (3) and (6), while the second, empirical approach functioned best with the model form based on Eq. (8). Although the derived patterns of vertical dispersion varied between the two sites, we used a single diurnal pattern for all years. We will later discuss our efforts to develop independent means of estimating the interannual variability of the pattern.

4. The regression models

The regression models shown in Table 1 were developed from an interpretation of the form of the simplified physical models described earlier. We have assumed that the statistical relationship between the different variables is of the form given by the physical models. We make no claims that the values of the fitted exponents of the regression models have particular significance. The simplest, model A, uses wind speed alone as a surrogate for all forms of dispersion and advection and is similar to a model used by Tiao et al. (1982) on CO data for Oregon. Model B corresponds to Eq. (3), model C to Eq. (6), and model E to Eq. (8). For comparison with model E, model D is included as a more conventional approach to incorporating a time-lag effect. In the models, TC_t is traffic count; WS_t, wind speed; and D_t, vertical dispersion, all at time t. A subscript of $t-1$ indicates a measurement from the preceding hour. The b, k, α, β and γ are fitted parameters, and e_t is the residual error term. In each model, the coefficient k is intended to be proportional to the CO emissions per vehicle near the CO monitor. Thus, the usefulness of a model will depend on the validity of k as a measure of emissions and on the precision of its estimates of k.

In developing model E, it was necessary to simplify Eq. (8) in order to obtain convergence in the nonlinear regressions. The exponential term of (8) has been incorporated into the estimated parameter, γ, modeling the time-dependent variation of the transport factor as a constant. This forces the actual time variation in that factor (which is not large in relative terms) to be incorporated in the residual term, e_t. The initial conditions, $C_0 h_0$, have been replaced by their steady state approximation for the previous hour, which is proportional to TC_{t-1}/WS_{t-1}. We found that the change in sign in the lag term was adequately captured by using this approximation.

5. Determining the best model

As an exploratory first step, models A–E were fitted to data for the separate years 1976, 1977 and 1978 using nonlinear least-squares regression. Each dataset contained 23 observations consisting of seasonal average data for each hour of the day excluding midnight. The need for one lag term meant that models D and E had $n = 22$. By fitting each year and season separately, we could observe whether the models performed consistently under different meteorological conditions.

Root-mean-square (rms) errors are shown in Table 2 for each of the models on CAMP data. Standard deviations of the original hourly average CO concen-

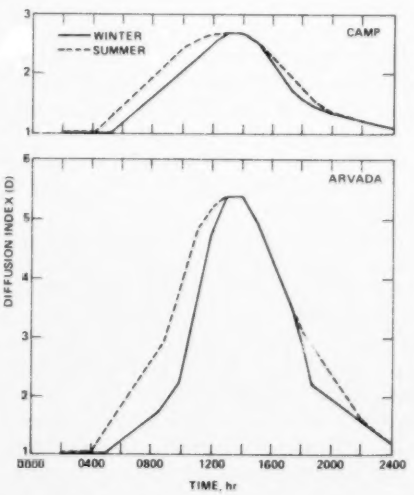


FIG. 3. As in Fig. 2 but for each monitoring site used with final regression model.

TABLE 1. List of models discussed.

Name	Model
Model A	$CO_t = b + \frac{k TC_t}{WS_t^{\alpha}} + e_t$
Model B	$CO_t = b\,\frac{k TC_t}{WS_t^{\alpha} D_t}\beta + e_t$
Model C	$CO_t = b + \frac{k TC_t}{WS_t^{\alpha} + D_t^{\beta}} + e_t$
Model D	$CO_t = b + \frac{k\left\{TC + \frac{\gamma TC_{t-1}}{WS_{t-1}^{\alpha} + D_{t-1}^{\beta}}\right\}}{WS_t^{\alpha} + D_t^{\beta}} + e_t$
Model E	$CO_t = b + \frac{k TC_t}{D_t^{\beta} WS_t}\left\{1 + \frac{\gamma}{WS_t}\left(\frac{TC_{t-1}/WS_{t-1}}{TC_t/WS_t} - 1\right)\right\} + e_t$

trations are shown in the first column for comparison. The rms error and standard deviation provide more information than a correlation coefficient by giving both the original variability in the data and the variability remaining after fitting the model. Reductions in the rms errors indicate improvements in model fit.

Note that the standard deviations are much higher for weekday CO data than for weekend data, and somewhat higher in winter than in summer. In terms of *proportion* of variance explained, the models are least effective on weekends when the variability in both CO and traffic counts is low. However, comparison of the models produces similar conclusions for all of the datasets. Clearly, both models B and C, incorporating vertical dispersion, offer considerable improvements over the fit achieved with model A. On practical grounds, too, model A was inadequate because its estimates of the "emissions" coefficient k were not significantly different from zero in most of the Arvada datasets.

TABLE 2. Root-mean-square errors for models A–E on CAMP data for 1976, 1977, and 1978.

CO dataset	SD of data	Model A	B	C	D	E
1976 winter						
weekday	2.77	1.50	0.72	0.75	0.62	0.69
weekend	0.87	0.57	0.29	0.32	0.32	0.38
1976 summer						
weekday	1.82	0.77	0.36	0.36	0.35	0.38
1977 winter						
weekday	2.92	1.46	0.83	0.86	0.46	0.32
weekend	0.97	0.45	0.19	0.19	0.18	0.29
1977 summer						
weekday	1.81	0.77	0.44	0.42	0.38	0.36
1978 winter						
weekday	2.69	1.35	0.90	0.91	0.72	0.41
weekend	0.76	0.64	0.35	0.41	0.24	0.19
1978 summer						
weekday	1.36	0.86	0.41	0.40	0.40	0.40

The poor performance of model A on Denver data, after its successful use for Oregon, is probably explained by the large amount of cloud cover routinely observed over Eugene, Oregon, masking the influence of the heat flux cycle. The average cloud cover for Eugene is greater than 8/10. Denver has a strong heat flux cycle, leading to differences in the diurnal cycles of wind speed and solar heating. The average cloud cover for Denver is 5/10 and wind speed in Denver peaks about 2 h later in the day than does the heat flux. In the absence of overcast, we expect vertical dispersion to follow heat flux. The failure of model A and other models of similar form, some including lag terms, provided the impetus for development of models containing estimates of vertical dispersion. Some of those models are described in Downton and Dennis (1983).

There was little difference between the fit of models B and C, indicating that the difference between the box-model equation (Eq. 3) and the scale-analysis equation (Eq. 6) may have little practical significance in this application. Models D and E, containing lag terms, noticeably improved the fit over model C in three datasets (1977 winter weekdays, 1978 winter weekdays and weekends). Judging from Table 2, the fit of models D and E is about equally as good.

The next step in comparing the models required investigating changes in CO emissions over a period of years. To estimate annual changes in CO emissions over the 1976–80 test period, the parameter k was allowed to vary by year while for all other parameters a single estimate across all years was made. To do that, in each model k was replaced by $k_y I_y$ where $I_y = 1$ in year y and 0 in other years. Models B to E were then

TABLE 3a. Model D at CAMP: parameter estimates for 1976–80 data (N = 110).

	Winter weekdays			Winter weekends			Summer weekday		
Parameter	Estimate	Standard error	Ratio	Estimate	Standard error	Ratio	Estimate	Standard error	Ratio
k_1	0.0169	0.0011	15.4	0.0253	0.0058	4.4	0.0154	0.0013	11.8
k_2	0.0151	0.0010	15.1	0.0150	0.0037	4.1	0.0085	0.0008	10.6
k_3	0.0146	0.0010	14.6	0.0135	0.0035	3.9	0.0123	0.0010	12.3
k_4	0.0157	0.0010	15.7	0.0179	0.0042	4.3	0.0101	0.0009	11.2
k^5	0.0138	0.0009	15.3	0.0195	0.0046	4.2	0.0091	0.0008	11.4
b	0.33	0.12	2.8	0.72	0.16	4.5	0.87	0.14	6.2
γ	0.66	0.25	2.6	3.06	1.04	2.9	0.78	0.29	2.7
α	0.68	0.05	13.6	0.89	0.09	9.9	0.35	0.15	2.3
β	0.74	0.03	24.7	0.80	0.08	10.0	0.78	0.03	26.0
rms error	0.61			0.44			0.59		
R^2	0.96			0.89			0.90		

fit to all 5 years of data for each of the six datasets. Again, models D and E provided the best fit as measured by R^2 and the rms errors.

The resulting parameter estimates for models D and E at both CAMP and Arvada are shown in Table 3. In each dataset, n = 110. The estimates, α, β, and γ serve as scaling parameters which affect the magnitudes of the k_y for each dataset. Within a dataset, relative changes in the k_y should show changes in emissions from one year to the next.

There is little difference in goodness of fit between the two models, judging from the R^2 values and rms errors in Table 3. However, the apparent similarity of fit, as measured by R^2, can be misleading because the independent variables are to some degree correlated. With correlated variables, parameters can easily switch greatly in magnitude from one dataset to another, even though there apparently is an excellent fit by the model. If the parameter values are switching around, then they may not be fully representing what one expects, because they are probably entangled with other parameters through the correlation. We note in Table 3 that for Arvada, the values of the emissions parameter, k_y, for model D are greatly different between winter weekdays and winter weekends, whereas, for model E they are basically the same. For CAMP the emissions parameters are basically the same for winter weekdays and weekends for both models. One does not expect a substantial change in emissions per vehicle between weekdays and weekends. This difference between model D and model E leads one to believe model E is the superior model; i.e., it is more correctly separating the effects of emissions from those of meteorology.

The models also differ in the precision of their estimates of the emission parameters k_y, which is important for the analysis of emissions change. The ratios of the estimates of the k_y to their standard errors show that the precision in the estimates of the k_y from model E is much higher than that from model D in five of the six datasets. (The exception is CAMP summer weekdays where the precision is approximately the same for the two models.) So model E appears to be better able to detect statistically significant changes in emissions. For these reasons, the next section of this

TABLE 3b. Model E at CAMP: parameter estimates for 1976–80 data (N = 110).

	Winter weekdays			Winter weekends			Summer weekday		
Parameter	Estimate	Standard error	Ratio	Estimate	Standard error	Ratio	Estimate	Standard error	Ratio
k_1	0.0382	0.0015	25.5	0.0503	0.0020	25.2	0.0323	0.0026	12.4
k_2	0.0352	0.0015	23.5	0.0325	0.0024	13.5	0.0169	0.0018	9.4
k_3	0.0348	0.0016	24.0	0.0301	0.0026	11.6	0.0251	0.0022	11.4
k_4	0.0336	0.0014	24.0	0.0364	0.0021	17.3	0.0214	0.0021	10.2
k^5	0.0303	0.0013	23.3	0.0395	0.0021	18.8	0.0195	0.0020	9.8
b	0.89	0.13	6.8	0.36	0.16	2.3	0.89	0.15	5.9
γ	1.33	0.07	19.0	0.72	0.24	3.0	1.08	0.10	10.8
β	0.96	0.05	19.2	0.86	0.05	17.2	1.14	0.09	12.7
rms error	0.62			0.45			0.64		
R^2	0.95			0.88			0.88		

TABLE 3c. Model D at Arvada: parameter estimates for 1976–80 data ($N = 110$)

	Winter weekdays			Winter weekends			Summer weekday		
Parameter	Estimate	Standard error	Ratio	Estimate	Standard error	Ratio	Estimate	Standard error	Ratio
k_1	0.374	0.057	6.6	0.0293	0.0076	3.9	0.0231	0.0018	12.8
k_2	0.319	0.050	6.4	0.0176	0.0047	3.7	0.0124	0.0010	12.4
k_3	0.243	0.037	6.6	0.0182	0.0048	3.8	0.0195	0.0015	13.0
k_4	0.155	0.022	7.0	0.0127	0.0033	3.8	0.0161	0.0013	12.4
k^5	0.227	0.033	6.9	0.0155	0.0040	3.9	0.0137	0.0011	12.5
b	0.91	0.09	10.1	1.14	0.07	16.3	0.54	0.04	13.5
γ	11.29	2.76	4.1	31.90	7.33	4.4	3.17	0.63	5.0
α	2.79	0.09	31.0	1.74	0.08	21.8	1.63	0.05	32.6
β	2.80	0.13	21.5	1.81	0.12	15.1	1.76	0.06	29.3
rms error	0.45			0.29			0.18		
R^2	0.95			0.90			0.94		

article focuses on an evaluation of the ability of model E to accurately estimate emissions change.

6. Analysis of emissions change

Although these models are ultimately expected to be used to estimate the effect of an inspection and maintenance (IM) program, the 1976–80 Denver data used here does not include the effects of IM. A gradual reduction in average emissions per vehicle is expected, because new cars with stricter emission standards are continually becoming a larger fraction of the vehicle fleet. Estimating this trend in emissions due to national standards will provide an evaluation of the model's ability to reliably estimate the effect of an inspection and maintenance program.

At both CAMP and Arvada, models D and E show fluctuation in the estimated k_y over the 5 yr, with an appearance of downward trend. To examine the validity of the best model, model E, we will cross-compare its emissions change estimates from the six datasets and then compare these estimates with an independent estimate based on direct measurement of emissions from automobiles.

7. Validity of regression model E

To examine the validity of the emissions change estimates obtained from model E, one would like to know, for some base period, what the emission changes in Denver actually were. The best estimate of actual emissions in past years can be obtained from the Environmental Protection Agency's mobile source emissions model, MOBILE2.5 (EPA, 1981). The mobile source emissions model is a computer program which estimates vehicle fleet emissions rates for hydrocarbons, carbon monoxide and nitrous oxides for any calendar year. The EPA believes there is sufficient information to establish realistic emission factors for these pollutants.

The MOBILE2.5 is based on direct measurements of emissions from vehicles of different model types, engine types, emissions control systems and ages. It accounts for vehicle type (cars, trucks, diesel cars, etc.), the degree of implementation of federal emissions

TABLE 3d. Model E at Arvada: parameter estimates for 1976–80 data ($N = 110$).

	Winter weekdays			Winter weekends			Summer weekday		
Parameter	Estimate	Standard error	Ratio	Estimate	Standard error	Ratio	Estimate	Standard Error	Ratio
k_1	0.0547	0.0020	27.4	0.0488	0.0023	21.2	0.0259	0.0009	28.8
k_2	0.0501	0.0019	26.4	0.0354	0.0021	16.9	0.0155	0.0007	22.1
k_3	0.0384	0.0017	22.6	0.0336	0.0020	16.8	0.0224	0.0008	28.0
k_4	0.0383	0.0014	27.4	0.0319	0.0016	19.9	0.0189	0.0007	27.0
k^5	0.0455	0.0016	28.4	0.0353	0.0018	19.6	0.0164	0.0007	23.4
b	0.36	0.09	4.0	0.78	0.08	9.8	0.27	0.04	6.8
γ	0.36	0.14	2.6	1.33	0.24	5.5	0.49	0.08	6.1
β	1.23	0.04	30.8	1.27	0.04	31.8	1.20	0.03	40.0
rms error	0.43			0.32			0.16		
R^2	0.95			0.88			0.95		

standards in each year, empirical estimates of actual emissions from vehicles (actual emissions are often far from the standards), empirical estimates of increases in emissions per mile as cars get older, differences between high- and low-altitude emissions, changes in average number of miles driven as a car gets older, and effects of ambient temperature, vehicle speed and operating mode (fraction of the miles driven under conditions of cold start, hot start, and stabilized operation). The MOBILE2.5 is essentially a "population" model of the fleet, with engineering parameters that can be defined to approximate the actual situation of vehicle operation

For this analysis, multiple runs of MOBILE2.5 under different parameter settings of temperature, hot-cold start fractions, average speed by road type, etc., were used to approximate average urban driving conditions at different times of the day, including rush-hour congestion. These calculations were specific to the vehicle mixes in Denver for two prototypic weekdays, one for winter and one for summer. Estimates of fleetwide average CO emissions, in grams per mile, were computed for each year from 1976 to 1981. Although there may be errors in the estimates of fleetwide emissions, because for example the relations between emissions and speed and between the test driving cycle and real driving are imperfectly known, we are fortunate to be dealing with cases for which traffic counts are essentially constant. This means the errors just mentioned become systematic ones which tend to be transparent when only relative changes in emissions over time are examined as done here. In terms of change over time the MOBILE2.5 estimates show a linear trend for 1976–80, with average CO emissions decreasing by 4.5% per year in winter and 6.2% per year in summer (see Fig. 4, top).

To estimate a corresponding linear trend using model E, the five emissions parameters k_y were replaced by the linear form

$$k_y = a(1 + my) \tag{9}$$

where $y = 0, 1, 2, 3, 4$ for the years 1976–80, respectively, and a and m are the parameters to be estimated. The resulting estimates of the slope m, expressed as a proportion of 1976 emissions, together with their standard errors, are shown in Table 4. All of the slope estimates are significantly different from zero at a 95% confidence level, and the weekday estimates are remarkably close to those computed by MOBILE2.5. This provides evidence of the validity of model E, i.e., that it measures what it was intended to measure (reliably separating the effect of emissions change from those of meteorology and traffic), when the emissions trend is linear. Also, differences between the slopes estimated for CAMP and Arvada are not statistically significant, indicating that model E is reliable over different datasets.

• *Comparison of regression and time series models.* A simpler method of modeling changes in ambient CO involves fitting a time series model to monthly averaged CO concentrations. Such a model was used by Tiao et al. (1982) in evaluating the Portland IM program. By using monthly averages of the data, the complex diurnal pattern of dispersion is ignored. The seasonal cycle of meteorological variation is dealt with by subtracting long-term monthly means from the data. Thus the data entered into the model are deviations from an average seasonal pattern. However, if a variable has no year-to-year variation, as is the case for our theoretical estimate of vertical dispersion, then that variable will drop out of the time series analysis. If such a time series model (without considering vertical dispersion) could perform as well as the regression models described earlier, then it would be preferable because the development of a theoretical or empirical dispersion pattern would not be necessary.

To compare the regression and time series approaches, a time series model was fit to the monthly average of the Denver traffic, and the wind speed data for 1976–81. Based on the physical models discussed earlier, models were considered of the form

$$\mathrm{CO}_t \propto \frac{E(\mathrm{TC}_t)}{\mathrm{WS}_t^{\alpha}} \tag{10}$$

where TC_t and WS_t are traffic count and wind speed in month t, respectively, and E is proportional to emissions per car passing the monitor. By taking logarithms of the data, this multiplicative model becomes additive. After taking logarithms, the effect of the annual weather cycle was removed by subtracting the mean value of each variable for the appropriate month from each data point.

Assuming a linear trend in emissions, the resulting model is

$$C_t = A + mt + T_t + \alpha W_t + e_t \tag{11}$$

where C, T, and W represent CO, traffic count, and wind speed for month t, respectively (in the form of log deviations); A, m, and α are fitted parameters, with m representing the proportional monthly change in emissions per vehicle. The residual errors, e_t, were found to be autocorrelated. Analysis of plots of the autocorrelation function and partial autocorrelation function showed that the best model of the errors was the first-order autoregressive model

$$e_t = \phi e_{t-1} + a_t \tag{12}$$

where ϕ is a fitted parameter and a_t is uncorrelated white noise. (The model fit and residual analysis were performed using the Statistical Analysis System (SAS) Autoreg procedure.)

The estimates of the relative emissions change parameter m for CAMP and Arvada, converted to an annual percentage change, are shown in Table 4. They are quite similar to the estimates obtained from

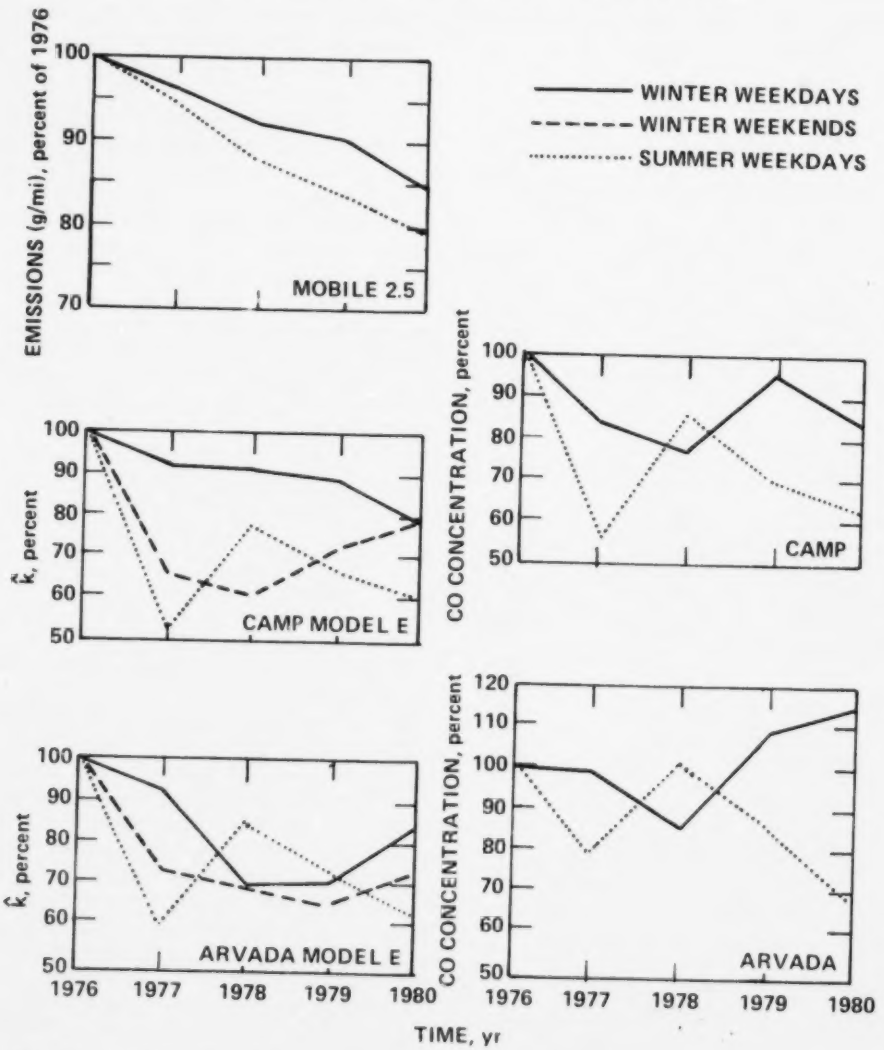


FIG. 4. Carbon monoxide emissions trends relative to 1976 (K); estimated by model E and by the mobil source emissions models, MOBILE2.5. Trends in ambient concentrations are also shown for comparison.

MOBILE2.5 and regression model E, considering that m represents an average of winter and summer emissions changes. However, the standard errors are larger than those of the regression model, with the result that the emissions change estimated for CAMP is not significantly different from zero at the 95% confidence level.

The smaller standard errors obtained with the regression model indicate that it can estimate emissions change with greater precision than the time series model. That enhanced precision, as we shall see in section 8, is crucial to distinguishing different trends in emissions produced by two separate causes: federal standards and IM. We conclude that the extra effort entailed in developing the regression model is worthwhile.

8. Sensitivity of regression model to detecting an IM effect

An important intended use of the regression model is to detect a change in the trend of CO emissions due to the institution of an IM program. How small can that change be and still be identified at some predefined level of significance? Based on our present work on Denver data, we can develop a simple first estimate of the size of change that can be detected. The precision

TABLE 4. Estimated annual change in vehicular emissions for the three types of models.

	MOBILE2.5 (1976–80)	
	Winter weekdays	Summer weekdays
Annual change:	−4.5%	−6.2%

	Regression model E, 1976–80 (n = 110)		
	Winter weekdays	Winter weekends	Summer weekdays
CAMP	−4.8%	−3.6%	−7.6%
(s.e.)*	(0.6)	(1.2)	(1.5)
R^2	0.95	0.69	0.75
Arvada	−5.5%	−5.4%	−6.0%
(s.e.)	(0.9)	(1.1)	(1.2)
R^2	0.91	0.81	0.82

	Regression model D, 1976–80 (n = 110)		
	Winter weekdays	Winter weekends	Summer weekdays
CAMP	−3.8%	−4.1%	−8.3%
(s.e.)	(0.7)	(1.2)	(1.2)
R^2	0.95	0.69	0.79
Arvada	−11.2%	−9.2%	−6.4%
(s.e.)	(1.1)	(1.4)	(1.4)
R^2	0.86	0.80	0.82

	Time series model, 1976–81 (n = 72)
	Full year
CAMP	−4.7%
(s.e.)	(2.7)
Arvada	−6.3%
(s.e.)	(1.6)

* Standard error.

of the estimate of the trend in emissions will play a key role.

To detect an effect of IM it will be necessary to distinguish a downward emissions trend due to IM from that due to the federal emissions control program. We can use EPA's mobile source emissions model for guidance, concentrating on the years 1975–90. That model estimates that the trend in emissions due to the federal emissions control program will be linear between 1975 and 1990. (The trend line is expected to flatten out to a constant by the late 1990s.) Implementation of an IM program several years before 1990 is expected to result in a steeper linear decrease in emissions until about 1988 and then parallel the non-IM trend line (the federal program alone). Thus, for an IM program instituted before about 1984, we can couch the detection problem in terms of detecting the difference in slope between two contiguous straight lines. For the Denver case, in which the IM program began in 1982, the best, most recent estimates for the post-IM emissions trend range between −6.3% and −6.9% per year (Colorado Department of Health, personal communication).

To estimate the smallest change in slope that can be detected by regression model E we will assume that we have the same number of years of post-IM data as used in fitting the model E. We will also assume that the residuals are independent, distributed normally, and that the magnitude of their variance is influenced solely by meteorology, not by the presence of an IM program. We are assuming that we obtain the same precision for the post-IM years as we have obtained for the pre-IM years. This latter assumption makes our conclusions optimistic. We then combine variances and use a t-test to test for differences in slope magnitude. Our test is

$$\frac{m_2 - m_1}{\sqrt{2}\,\text{s.e.}} \geqslant \text{critical value of Student's t}$$

where s.e. is standard error.

In selecting a confidence level for determining statistical significance, it is common practice to try to minimize the type I error to avoid erroneously rejecting a hypothesis which is actually true. Based on that criteria a confidence level of 95% or 99% is commonly selected. (In this application, the null hypothesis would be that there has been zero change due to IM. A type I error would be finding that there was a change due to IM and continuing the program when it was, in fact, ineffective.) In this application, however, we are also seriously concerned about a type II error—erroneously accepting a hypothesis which is actually false. (In this example, a type II error would be finding no change due to IM and discontinuing the program when there had actually been a change due to IM that the test failed to pick up.) Erroneously halting an effective IM program may be as bad as erroneously keeping an ineffective IM program.

The relative seriousness of these two errors is a political decision, not a statistical one. Recognizing that a type II error is important, however, we recommend the choice of a 90% confidence level. We have chosen to use a one-tailed test because we are looking only for a decrease in CO. An increase due to IM is not considered plausible. The critical value of Student's t, then, is $t_\alpha = 1.28$.

For winter weekdays at CAMP and Arvada (see Table 4) the difference in slope must be greater than or equal to 1.1 and 1.6% per year, respectively, to be positively identified. For summer weekdays at CAMP and Arvada, the difference in slope must be greater than or equal to 2.7 and 2.2% per year, respectively. For Denver, which is at high altitude (5283 MSL), more recent

analyses with the newer model, MOBILE3 predict a "perfectly running program" to have a difference of slope ranging between 3.4% and 4.4% per year under the best of circumstances. Under more realistic assumptions, modifying MOBILE3 parameters to reflect more closely the implemented IM program, one expects the difference to be between 1.8% and 2.4%. The MOBILE 3 is used because for post-1981 cars, it contains the most up-to-date data on control technology type, emissions performance and amount of tampering, all of which affect estimates of IM program effectiveness (U.S. EPA, 1985).

Thus, it appears that a regression model similar to model E could detect a statistically significant change in emissions trend for a well-operating IM program for winter weekday data, but possibly not for summertime data. Based on MOBILE3, low-altitude programs seem to fall into the same range; the difference in slope for "perfect" programs is 3.5 and for more realistic estimates the change is 2.1.

9. Attempts to improve the regression model

In the regression models described heretofore, wind speeds, traffic counts, and CO concentrations varied from year to year but the diurnal pattern of vertical dispersion was not varied. There is a danger that actual year-to-year variations in dispersion, not accounted for in the models, could lead to biased estimates of emissions changes due to misrepresentation of the parameter k.

In the version of model E which contains a linear emission change, the residuals from the regression were clearly not white noise in five of the six datasets. In particular years, the residuals were either predominantly positive or predominantly negative. Referring back to the results of model E shown in Table 3, in which a separate k_y was estimated for each year, it can be seen that the k_y decrease smoothly only for CAMP winter weekdays. In the other datasets, the k_y fluctuate about the decreasing trend. Figure 4 shows the patterns of estimated emissions, with k_y for each year plotted as a percentage of the 1976 emissions estimate. For comparison; corresponding emissions estimates from MOBILE2.5 are also shown.

Model E closely reproduces the MOBILE2.5 estimates on winter weekdays at CAMP. However, the fluctuations in k_y in the other datasets indicate that those parameters are not serving purely as a measure of emissions per vehicle. It appears that some year-to-year variations in meteorology are not accounted for by wind speeds. Such variations would be incorporated in the estimates of k_y in the least-squares fitting process, thereby potentially biasing the "emissions" estimate. (Fortuitously, in our particular case the variations do not seem to have seriously biased the estimates, given the consistent good agreement evident in Table 4.) When similar fluctuations occur at the two monitoring sites we expect that region-wide meteorological factors are to blame. (See particularly the summer weekdays, where similar variations occur at both sites. A similar pattern of variability is found in the original summer average CO concentrations as well, as shown in Fig. 4.)

To learn more about what factors might be missing from model E, we ran the model "backwards". That is, we used the CO, traffic, and wind speed data, and the values of k_y predicted by MOBILE2.5 to "solve" for the dispersion function D that would give a perfect fit. The Arvada and CAMP results are shown in Fig. 5 for each year, separately for winter and summer. The curves are similar to measurements of diurnal variation of sensible heat flux or vertical turbulence, such as sigma-w, measured in other urban areas (e.g., see Ching et al., 1983, Clarke et al., 1982). This reinforces the earlier conclusion from theoretical arguments that vertical dispersion is the major missing variable.

Exploring the available data from Denver's Stapleton Airport, we examined routine meteorological measurements directly related to vertical dispersion, hoping to find data that could explain the interannual variations shown in Fig. 5. These curves show the most year-to-year variability in the middle of the day, during the time of major solar insolation. Lapse rates and estimated mixing heights from 0500 and 1700 LST soundings were investigated but showed little promise of explaining midday variability.

The most promising approach resulted from estimating the Monin-Obukov length, L, from surface meteorological data using the method of Holtslag (Holtslag, 1984 and Paumier et al., 1986). The length L is an indicator of whether combined effects of thermal stratification and mechanical mixing tend to enhance or suppress turbulent fluctuations in the vertical. The midday pattern of $\exp(-10/L)$ approximated our extracted vertical dispersion function D, correctly duplicating the time of the peak in summer and winter. Variations in seasonally averaged peak values of $\exp(-10/L)$ matched those in Fig. 5 for Arvada for winter, but not for summer. The failure on summer data could be caused by the fact that some convective processes, most important in summer, are not distinguished in the computation of L. For example, localized wind circulations resulting from thunderstorms, common over Denver in the summer, can be very effective at dispersing material.

Use of the Monin-Obukov length in the regression model was tried only for Arvada, which has a suburban environment reasonably similar to that at Stapleton airport. Measurements take in Stapleton's suburban-rural environment were not expected to be representative of conditions at the downtown Denver CAMP site. For Arvada, the seasonally averaged peak values of $\exp(-10/L)$ were used to make annual adjustments in the dispersion functions D for several hours in the middle of the day. This resulted in very minor changes

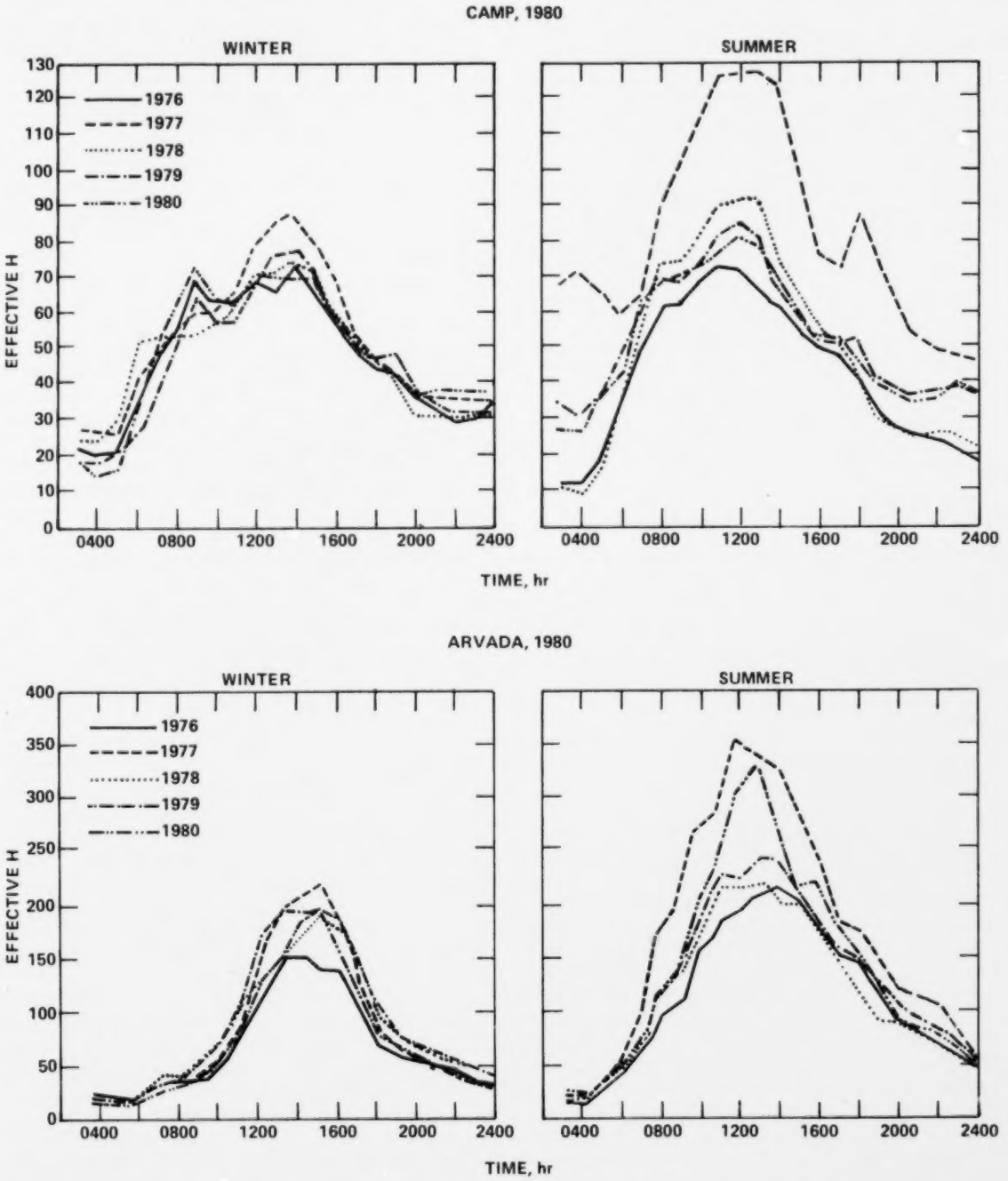


FIG. 5. Diurnal pattern of vertical diffusion (effective H) necessary to produce the most perfect fit to the carbon monoxide data for each season with model E.

in the estimates of the k_y. Although the changes were in the right direction, they were much too small to be considered a real improvement.

Since the effective mixing height can be affected by more than just heating, we explored a number of "second-order" effects to see if they could explain the mid-

day interannual variation of vertical mixing. We examined monthly averaged wind direction, rainfall, cloudiness, heating degree days (as an indirect measure of insolation), and cooling degree days. None of these variables had year-to-year variations at all similar to the variation of the inferred vertical mixing.

Our attempt to find explanations for the apparent interannual variability in dispersion by developing a single measure from routinely monitored meteorological data gave discouraging results. It appears that a more direct measure of vertical mixing near the surface is needed, preferably at a location close to the CO monitor. We would suggest that measurement of sigma-w might meet the need for a direct measure. It is proportional to the convective velocity scale and also to heat flux; it should be a valid measure at midday for both summer and winter, and its diurnal pattern is similar to the curves in Fig. 5.

10. Conclusions

From the above analysis we derive four conclusions: 1) It is important to have both wind speed and vertical dispersion represented in regression models used to fit diurnal patterns of pollutant concentrations. 2) We found a strong concurrence between theoretical arguments concerning model form and our results of fits to the data for different regression model forms. 3) The difference in the diurnal shape of wind speed versus vertical dispersion appears to be very important. We expect this conclusion to be valid for other urban areas with moderate to strong insolation. In addition, 4) our new approach to modeling of time lags in regression models, allowing their influence to be both positive and negative, gave very good results in terms of model precision.

The regression model we developed verifies well with independent information. The emissions trends estimated for weekday cases are remarkably close to those computed by the mobile source emissions model based on direct measurement of emissions from automobiles. Differences between trends estimated for the two monitoring sites are not statistically significant, indicating that the regression model is reliable over different datasets. Both results indicate that the model is capable of reliably separating the effects of emissions change from fluctuations in meteorology and traffic.

The regression model we developed has sufficient precision to identify present trends in emissions from ambient concentration data. Also, it appears that for winter meteorology, this model may have sufficient precision to detect changes in an emissions trend due to implementation of an IM program. The precision may or may not be adequate to detect changes for summer meteorology.

We were not able to explain the interannual variability in vertical dispersion using routinely available surface meteorological data. Examination of standard errors indicates there is still room for improvement of model precision by accounting for this interannual variability and by better accounting for summertime meteorology. A logical candidate to provide the needed new information for more advanced regression analyses is measurement of sigma-w. The Colorado Department of Health has begun to collect such data at the two monitoring sites used in this study. Thus it would be profitable to revisit our analysis when a sufficient amount of that data (a few years') becomes available.

REFERENCES

Anthes, R. A., 1979: Meteorological aspects of regional-scale air-quality modeling. *Advances in Environmental Science and Engineering*, 1. Pfafflin and Ziegler, Eds., Gordon & Breach, 3–49.

Carson, D. J., 1973: The development of a dry inversion-capped convectively unstable boundary layer. *Quart. J. Roy. Meteor. Soc.*, **99**, 450–467.

Ching, J. K. S., J. F. Clarke and J. M. Godowitch, 1983: Temporal behavior of turbulence parameters relevant to elevated point source dispersion. *Sixth Symposium on Turbulence and Diffusion*. Amer. Meteor. Soc.

Clarke, J. F., J. K. S. Ching and J. M. Godowitch, 1982: An experimental study of turbulence in an urban environment. EPA Technical Report, EPA-600/3-82-062, Research Triangle Park, NC, 150 pp.

Downton, M., and R. Dennis, 1983: Modeling ambient carbon monoxide trends to evaluate automobile emission controls. *Eighth Conference on Probability and Statistics in Atmospheric Sciences*. Amer. Meteor. Soc., 212–219.

Holtslag, A. A. M., 1984: Estimates of diabatic wind speed profiles from near-surface weather observations. *Bound.-Lay. Meteor.*, **29**, 225–250.

Ledolter, J., and G. C. Tiao, 1979: Statistical models for ambient air pollutants, with special reference to the Los Angeles Catalyst Study. *Environ. Sci. Tech.*, **13**, 1233–1240.

Lettau, H., 1970: Physical and meteorological basis for mathematical models of urban diffusion. *Proc. of Symposium on Multiple-Source Urban Diffusion Models*, U.S. EPA, APCO Publ. No. AP-86.

Long, G. E., P. Middleton, R. L. Dennis and R. W. Miksad, 1982: An urban air quality box model for carbon monoxide. *Eos*, **63**, 894.

——, ——, —— and ——, 1983: Carbon monoxide modeling for urban air quality management. *Intern. J. Environ. Studies*, **21**, 215–228.

Paumier, J., D. Stinson, T. Kelly, C. Bollinger and J. S. Irwin, Feb. 1986: MPDA-1: A meteorological processor for diffusion analysis, User's Guide. U.S. EPA Report, EPA-600/8-86-011, Research Triangle Park, NC, 182 pp.

Tennekes, H., 1973: A model for the dynamics of the inversion above a convective boundary layer. *J. Atmos. Sci.*, **30**, 558–567.

Tiao, G. C., G. E. P. Box and W. J. Hamming, 1975: A statistical analysis of the Los Angeles ambient carbon monoxide data 1955–1972. *APCA Journal*, **25**, 1129–1136.

——, J. Ledolter and G. B. Hudak, 1982: Statistical analysis of the effect of car inspection and maintenance programs on the ambient CO concentrations in Oregon. *Environ. Sci Tech.*, **16**, 328–335.

U.S. EPA, Ann Arbor, Office of Mobile Source air pollution control, Feb. 1981: Users Guide to MOBILE2 (Mobile Source Emissions Model). EPA Report 460/3-81-006, Springfield, VA.

——, Ann Arbor, Office of Mobile Sources, September 1985: Compilation of air pollutant emission factors, Volume II: Mobile Sources, EPA Report AP-42. Fourth Edition, Ann Arbor, MI, 555 pp.

NOTICE

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

This article documents research that was conducted under the auspices of the U.S. Environmental Protection Agency (USEPA). As an arm of the government, the USEPA has a paid up, non-exclusive, irrevocable, worldwide license to reproduce, prepare derivative works, and distribute to the public all of the material contained herein, and to grant others acting on its behalf the right to do so. The USEPA grants to the National Technical Information Service, Department of Commerce, the right to act as its agent in the distribution of this material.

1391.a